working paper
department
of economics

BASINS OF ATTRACTION, LONG RUN EQUILIBRIA, AND THE SPEED OF STEP-BY-STEP EVOLUTION

Glenn Ellison

96-4 Aug. 1995

massachusetts
institute of
technology

50 memorial drive
cambridge, mass. 02139

BASINS OF ATTRACTION, LONG RUN EQUILIBRIA, AND THE SPEED OF STEP-BY-STEP EVOLUTION

Glenn Ellison

96-4 **Aug. 1995**

Basins of Attraction, Long Run Equilibria, and the Speed of Step-by-Step Evolution

Glenn Ellison[1]

Department of Economics
Massachusetts Institute of Technology

August 1995

[1]I would like to thank Sara Fisher Ellison, Drew Fudenberg, David Levine, Eric Maskin, Roger Myerson, Georg Nöldeke, Tomas Sjöström and Peter Sorensen for their comments. Financial support was provided by National Science Foundation grants SBR-9310009 and SBR-9496301.

Abstract

The paper provides a general analysis of the types of models with ϵ-perturbations which have been used recently to discuss the evolution of social conventions. Two new measures of the size and structure of the basins of attraction of dynamic systems, the radius and coradius, are introduced in order to bound the speed with which evolution occurs. The main theorem uses these measures to provide a characterization useful for determining long run equilibria and rates of convergence. Evolutionary forces are most powerful when evolution may proceed via small steps through a series of intermediate steady states. A number of applications are discussed. The selection of the risk dominant equilibrium in 2×2 games is generalized to the selection of $\frac{1}{2}$-dominant equilibria in arbitrary games. Other applications involve two-dimensional local interaction and cycles as long run equilibria.

Keywords: Learning, long run equilibrium, equilibrium selection, waiting times, radius, coradius, risk dominance, local interaction, Markov process.

JEL Classification No.: C7

Glenn Ellison
Department of Economics
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139

1 Introduction

Multiple equilibria are common in economic models, and when they are present, both aggregate welfare and the distribution of payoffs often varies dramatically with the equilibrium on which players coordinate. Beginning with the work of Foster and Young (1990), Kandori, Mailath, and Rob (KMR) (1993), and Young (1993a), a large number of recent papers have explored the extent to which evolutionary arguments can provide a rationale for regarding some equilibria as being more likely than others. This paper attempts to contribute to this literature in two ways. First, it develops a general technique for analyzing both the long run and the medium run behavior of such models, providing intuition for why the results of previous analyses are what they are, and allowing for more thorough descriptions of behavior. Second, it discusses a number of specific topics which should be of independent interest. Most prominent among these (and hopefully of broader interest outside the literature on the development of conventions in games) is a discussion of when evolutionary forces produce rapid change (and hence are practically relevant), in which it is argued that a crucial determinant of whether evolutionary change will occur rapidly is whether it is possible for change to be effected gradually via a number of small steps.

The papers of KMR, and Young (1993a) develop explicit models of processes by which conventions might be established as large populations of boundedly rational players interact and adjust their play over time. Each makes the assumption that there is some source of persistent noise in the population, *e.g.* player turnover, trembles, etc., with the central observation being that small amounts of noise may dramatically alter the long run behavior of the system. While any behavior is in theory possible once noise is introduced, some actions become much more likely than others, which suggests the definition of "long run equilibria" as states which continue to be played in the long run with nonvanishing probability as the amount of noise tends to zero. The most striking result which these papers obtain is that such an analysis not only rules out unstable mixed equilibria, but usually provides an argument for selecting an unique equilibrium even in games with multiple strict Nash equilibria. For the case of 2×2 games, these papers find that for a variety of dynamics the unique long run equilibrium involves the population coordinating on the equilibrium Harsanyi and Selten (1988) term risk-dominant. A number of more recent papers have explored how this conclusion is affected by the assumptions about the nature of the social interactions

and learning rules, and what similar analyses have to say about bargaining, signalling, and other classes of games with multiple equilibria of inherent interest to economists.[1]

The first part of this paper attempts to develop a new framework for analyzing models with persistent noise. In previous papers, the standard approach has been to rely on a tree construction algorithm developed by KMR, Young (1993a), and Kandori and Rob (1992) (building on the work of Freidlin and Wentzell (1984)) to identify the long run equilibria. This method of analysis has a number of significant limitations which provide the motivation for this paper. The first obvious limitation of the technique is that the analysis it provides is limited in scope in addressing only long run behavior and not saying anything about how rapidly the long run equilibrium is reached.[2] As Ellison (1993) argues, convergence times for models in this literature are at times so incredibly long that long run behavior may have very little to do with what what happens within any economically relevant time horizon. Another drawback is that while the tree construction algorithm is in theory universally applicable, it has in practice proven difficult to apply to complicated models and when developing generalizations. Finally, the nontransparency of the algorithm has make the whole literature seem a bit mysterious — a common frustration is the feeling that the typical paper writes down a model, says something about trees and after several pages of calculations gives an answer but provides little understanding of how the answer is connected to the assumptions of the model.

The framework developed in this paper provides a method for analyzing both the long run and the medium run behavior of models with small persistent noise. While the method is not universally applicable, it provides an intuitive understanding of the literature to date and may facilitate both more thorough analyses and analyses of more general and complex models in the future.

One can think of the main theorem of the paper as incorporating three main insights. First, it is noted that virtually all of the models in the literature share a common structure

[1]See, for example, Bergin and Lipman (1993), Binmore and Samuelson(1993), Ellison (1993), Kandori and Rob (1992, 1993), Nöldeke and Samuelson (1993, 1994), Robson and Vega Redondo (1994), and Young (1993b).

[2]The papers of Ellison (1993), Binmore and Samuelson (1993), and Robson and Vega Redondo (1994) are notable exceptions to the tendency to rely excusively of the Freidlin-Wentzell construction and thereby ignore convergence rates.

and thus can be addressed within the context of a simple (albeit abstract) model. Next, the paper points out a relationship between the long run behavior of models with noise and the speed with which evolutionary change occurs which at least occasionally provides clear intuition for why the long run equilibrium of a model is what it is. The observation is simply that if a model contains a state or set of states Ω which is both very *persistent* in the sense that play tends to remain in a neighborhood of Ω for a long time whenver Ω is reached, and sufficiently *attractive* in the sense that Ω will be reached relatively quickly after the system begins in any other state, then in the long run states near Ω will occur most of the time. In light of this observation, an understanding of what makes evolution evolution fast or slow will provide also an algorithm for identifying long run equilibria.

The final crucial observation, which should be of interest to the broadest audience, concerns the speed with which evolutionary change occurs Specifically, it is argued that a critical determinant of whether evolutionary change will occur rapidly is whether the process may proceed by a (possibly very long) series of gradual changes from one nearly stable state to another, or whether more dramatic changes are required. A biological analogy helps one think about why evolutionary mechanisms of the former type are more plausible. Think of how a mouse might evolve into a bat. If the process of growing a wing required ten distinct independent genetic mutations and a creature with anything less than a full wing was not viable, we would have to wait a very, very, long time until one mouse happened to have all ten mutations simultaneously. If instead a creature with only one mutation was able to survive equally (or had an advantage, say, because a flap of skin on its arms helped it keep cool), and a second mutation at any subsequent date produced another viable species, and so on, then evolution might take place in a reasonable period of time. While this paper focusses on the development of conventions in games, I hope that the observations on the speed of evolution which this paper provides may be of use more generally in helping to assess arguments in a variety of areas of economics where "evolutionary" changes are discussed.

The main theorem of the paper combines and formalizes these observations a characterization of the long run and medium run behavior of models with noise. The formalization focusses on two new measures, referred to as the radius and modified coradius, which describe the size of the basin of attraction of a limit set and its relation to the basins of

attraction of the other limit sets — and which can be used to provide bounds on the speed with which evolution occurs. The radius measure, denoted by $R(\Omega)$, is obtained by simply counting the number of "mutations" needed to escape the basin of attraction. It provides an upper bound on the speed with which evolution away from Ω can occur, capturing the notion of persistence. The modified coradius measure, $CR^*(\Omega)$, provides a lower bound on the speed with which evolution toward Ω will occur. The measure itself is somewhat more complicated so as to capture the fact that Ω will arise more quickly both if it can be reached with few mutations and if it is possible to reach it via a number of smaller transitions between intermediate limit sets. Formalizing the intuition that a state which is both persistent and sufficiently attractive will occur most of the time, the main theorem of the paper shows that Ω is the unique long run equilibrium of a model if $R(\Omega) > CR^*(\Omega)$. In this case the modified coradius is also shown to provide a bound on how rapidly Ω will be reached.

One can see this theorem as making a number of contributions to the literature. First, the most obvious benefit is that it provides a tool with which one can analyze medium run as well as long run behavior. Second, what may be the most important and unexpected benefit is the intuition which the theorem provides for what is going on in the literature. While the reader will probably not be surprised to learn that long run equilibria can in some cases be identified by looking at simple measures of persistence and attractiveness, what I think is unexpected is that the main theorem of this paper is sufficiently powerful so as to allow virtually all of identifications of long run equilibria which have been given in the past to be rederived as Corollaries (albeit sometimes with a great deal of work). An implication of this broad applicability is that behind most of the mysterious tree constructions in the literature, all that is happening is that there is a nearly stable state (or a set of such states) which requires a large number of mutations to escape (and hence is persistent) and to which the system returns relatively quickly after starting at any other point (either because it can be reached with few mutations or by a sequence of smaller steps through intermediate limit sets). Finally, I hope that the theorem will prove valuable also in helping researchers to identify the long run equilibria of complex and general models. The examples and corollaries which are presented are intended in part to illustrate where (and how) the theorem may be most usefully applied.

Following the main theorem, the paper explores a number of different topics which should be of independent interest (at least to those who have followed the literature on the development of conventions in games.) The first of these involves generalizing the famous result that risk-dominant equilibria are selected in 2×2 games. The discussion here first reviews negative results (providing a new example which illustrates the nonrobustness of the long run equilibrium concept by showing how outside the class of 2×2 games the selected equilibrium may depend on which Darwinian behavior rule is chosen), and then presents the new positive result which is the main item of interest. The result is that in any game for which an equilibrium satisfying Morris, Rob and Shin's (1995) refinement of risk dominance, $\frac{1}{2}$-dominance, exists, that equilibrium is selected by the KMR model, and that this selection is quite robust in that it holds both for any Darwinian dynamic and in many local interaction settings. The trivial proof of this result serves also to illustrate how the main theorem may aid in the derivation of general results.

Two additional learning in games topics which are discussed at some length are local interaction models and the extent to which models of evolution with noise should be thought of as a method of equilibrium selection. The primary result on local interaction is that $\frac{1}{2}$-dominant equilibria are selected also in a two-dimensional local interaction model, with convergence to the long run equilibrium being relatively fast. This latter result is noteworthy because the model lacks the powerful "contagion" dynamics which one might have thought drove the fast evolution of one dimensional local interaction models. The model serves also to clarify my earlier comment that the techniques developed here are most likely to prove useful in analyzing complex models — the ability to limit one's focus to a single limit set (and paths to it) and the ability to account for the impact of transitions through intermediate limit sets are each most valuable when considering models for which the unperturbed dynamics feature a large number of steady states and cycles.

The main comment of this paper on the equilibrium selection interpretation of evolutionary models with noise, which is illustrated via a simple example, is that the term long run equilibrium may be something of a misnomer in that an evolutionary model can easily select a cycle or other behavior even in games with an unique Nash equilibrium.

The final two sections of the paper are concerned with its relationship to the literature. The first contains a discussion of the applicability of the main theorem which attempts to

catalog both the extent to which the results of previous papers could have been derived as applications of the main theorem (in which case it provides intuition and convergence rates) and the extent to which the main theorem would have allowed results to be derived much more easily. Using the former criterion the theorem is very widely applicable; the latter situation is less common. The final section contains an alternate proof of the long run equilibrium part of the main theorem based on a Freidlin-Wentzell "tree surgery" argument. In light of this argument one can interpret this part of the theorem also as illustrating how a systematic tree surgery argument can be applied to models in which transitions through intermediate limit sets are important. This section will clarify how this paper builds on the techniques of Young (1993a), Kandori and Rob (1992), Ellison (1993), Evans (1993), Samuelson (1994), and others, and suggests how "tree surgery" arguments might be applied in other contexts.

2 Model

One of the primary goals of this paper is to provide an improved understanding of the behavior of evolutionary models with noise. As a first step toward this goal, I note here that if one abstracts away from the details it becomes clear that virtually all of the models in the literature can be fit into a framework which is remarkably simple. Because the economic motivation for this abstract model would otherwise be a total mystery to anyone not familiar with the literature, I begin this section with a concrete example and then describe the general framework within which the main results of the paper will be formulated.

2.1 An example — a model of the evolutions of conventions in games

The primary motivation for the recent literature on models of evolution with noise is the suggestion of Foster and Young (1990), KMR and Young (1993a) that we might better understand the behavior of players in games (especially those with multiple equilibria) by explicitly modeling the process by which "social norms" or "conventions" might develop as boundedly rational players adjust to their environment over time. In this section I present one such example which is meant to illustrate the type of model which the abstract framework discussed later is intended to encompass.

Let G be a symmetric $m \times m$ game. We will suppose that conventions for how to play in

this game develop within a society of N players as they are repeatedly randomly matched to play G in periods $t = 0, 1, 2, \ldots$, with each player choosing a single action to use in each period not knowing who his opponent will be. We can describe the play of the population at time t by a vector $(s_1, s_2, \ldots, s_N)$ giving the strategies of each of the players. We will write Z for the set of all such strategy profiles, and refer to the elements of Z as the possible states of the system. It simplifies notation at times to pick an (arbitrary) ordering of the states so that each can be represented also by a $1 \times m^N$ vector, z_t, equal to one in a given place and zero in all the rest.

Rather than suppose that the players choose their actions rationally, one might imagine instead that the players react to their environment by following some simple boundedly rational rule. One particular example which has been extensively studied is the "best-response dynamic" under which each player at time t plays a best response to the strategies used at time $t-1$ by his potential opponents. We can incorporate this or any other deterministic behavior rule in which players react only to actions in the previous period by specifying that changes in play over time are described by the appropriate "deterministic evolutionary dynamic", a function $b : Z \longrightarrow Z$ such that $z_{t+1} = b(z_t)$.

When one is trying to model the behavior of a large society, an argument can certainly be made that it is not plausible to assume that the specified boundedly rational learning rule will describe everyone's play. More reasonably, we might assume instead that there will always be noise on top of this due to trembles in action choices, players reacting to idiosyncratic payoff shocks, the replacement of old players with new players unfamiliar with the prevailing norms, and other factors. The crucial observation of Canning (1992), KMR, and Young (1993a) is that while long run behavior an evolutionary model without noise will typically be dependent upon initial conditions (which we usually think of as a result of historical accidents we know little about), this dependence is eliminated if a sufficient amount of noise is introduced into the model. As a result, we may obtain a model in which we can more confidently predict the long run course of play.

The most common specification of noise in the previous literature is to suppose that for some fixed $\epsilon \in (0, \frac{1}{m})$, the evolution of play over time is described by a stochastic process which is defined by composing the deterministic dynamic b with a random mutation process under which each player's strategy at time $t+1$ is left unaltered with probability $1 - m\epsilon$,

and is drawn independently from a distribution placing equal probability on each of the possible strategies with probability $m\epsilon$. To describe such a model mathematically, note that given our notation with z_t as a vector, the deterministic dynamics may be represented as a linear transformation $z_{t+1} = z_t P$, with P an $m^N \times m^N$ matrix which has a one in the ij^{th} place if $b(i) = j$ and a zero there otherwise. The stochastic model is a Markov process on the state space Z. The probability distribution over the possible states at time t is described by a $1 \times m^N$ vector v_t whose i^{th} element gives the probability of the state z_t which has a one in that place. We have

$$v_{t+1} = v_t P(\epsilon),$$

where $P(\epsilon)$ is the probability transition matrix given by

$$P_{ij}(\epsilon) = \epsilon^{c(i,j)}(1 - (m-1)\epsilon)^{N-c(i,j)},$$

with $c(i,j)$ the number of players who play differently in the states j and $b(i)$. Note that $c(i,j)$ is a measure of the relative likelihood of transitions when ϵ is small (with high values of $c(i,j)$ corresponding to very unlikely transitions). It is standard to refer to $c(i,j)$ as the *cost* of a transition from i to j. I will later use the fact that this cost function is of the form $c(i,j) = d(b(i),j)$ with d satisfying the triangle inequality, *i.e.* $d(x,z) \leq d(x,y) + d(y,z)$.

2.2 The general model

I now discuss the more abstract model in the context of which the general characterizations of the behavior of evolutionary models of noise will be developed. The model abstracts away from many details of the evolutionary process in order to obtain a framework which is sufficiently general so as to encompass most of the models which have previously been studied. A nice side benefit of this generality is that the framework becomes quite simple and therefore helps clarify the structure of these models.

Essentially, (most) previous models of learning and evolution with noise can be seen as combining various specifications of three main elements of the general model defined below. The first main element of the model is a finite set Z which I will refer to as the *state space* of the model. In the typical model the state space consists of a set of possible descriptors of how a game is being played. For example, a state $z \in Z$ might be the current

strategy profile of the population, a summary statistic such as the number of players using each strategy, or a more detailed description of play in which the current strategy profile is augmented with historical information on past play or on who was matched with whom.

The second main element of the model is a Markov transition matrix P on the state space Z. In the typical model P describes the implications of a particular specification of the process by which boundedly rational players adjust their play over time. For example, besides the best-response dynamics it could be used to capture a variety of more complex decision rules where players make use of more information (for example historical data) or which are stochastic because players adjust their play in each period with some probability or rely on information the acquisition of which is affected by the outcome of the random matching process. Note that the one critical restriction here is that P is Markov, so that one must include in the state space all information on which players decisions are to be based. For example, if players are assumed to play a best response to their most recent k observations, the state should incorporate what each player saw in the previous k periods.

Finally, the third main element of the model is a specification of some type of small random perturbations. In particular, I will assume that for each ϵ belonging to some interval $[0, \bar{\epsilon})$, we are given a Markov process $\{z_t^\epsilon\}$ with state space Z and with transition probabilities

$$\text{Prob}\{z_{t+1}^\epsilon = z' | z_t^\epsilon = z\} = P_{zz'}(\epsilon)$$

such that (i) for each $\epsilon > 0$ the transition matrix $P(\epsilon)$ is ergodic; (ii) $P(\epsilon)$ is continuous in ϵ with $P(0) = P$; and (iii) there exists a cost function $c : Z \times Z \to \mathcal{R}^+ \cup \infty$ such that for all pairs of states $z, z' \in Z$, $\lim_{\epsilon \to 0} P_{zz'}(\epsilon)/\epsilon^{c(z,z')}$ exists and is strictly positive if $c(z, z') < \infty$ (with $P_{zz'} = 0$ for sufficiently small ϵ if $c(z, z') = \infty$). The characterizations we will provide of medium run and long run behavior depend on the specification of the random perturbations only through the cost function, so we will tend to think of the cost function as a primitive of the model concerning the relative likelihood of the possible mutations. For example, instead of simply choosing the cost of a transition to be the number of independent random trembles necessary for it to occur, a cost function could be chosen so as to incorporate state-dependent mutation rates with unbounded likelihood ratios (as in Bergin and Lipman (1993)), correlated mutations which occur among groups of players, or the impossibility of some mutations (in which case the cost is set to infinity).

2.3 Descriptions of long run and medium run behavior

In this paper long run behavior will be described in the manner which has become standard in the literature. Specifically, for any fixed $\epsilon > 0$, the Markov process corresponding to the model with ϵ-noise has an unique invariant distribution μ^ϵ (given by $\mu^\epsilon = \lim_{t\to\infty} v_0 P^t(\epsilon)$) which we think of as describing the expected likelihood of observing each of the states given that evolution has been going on for a long, long time. While it is this distribution which is of most interest, μ^ϵ is difficult to compute, and hence it has become standard in the literature to focus instead on the *limit distribution* μ^* defined by $\mu^* = \lim_{\epsilon\to 0} \mu^\epsilon$.[3] The motivation for this is that μ^* is easier to compute and provides an approximation to μ^ϵ when ϵ is small. Also following the literature we will call a state z a long run equilibrium if $\mu^*(z) > 0$.[4]

The most basic result on what long run equilibria look like (found in Young (1993a)) is that the set of long run equilibria of the model will be contained in the recurrent classes of $(Z, P(0))$.[5] The recurrent classes are often referred to as *limit sets*, because they indicate the sets of population configurations which can persist in the long run absent mutations. They may take the form of steady states, of deterministic cycles, or of a collection of states between which the system shifts randomly. The prototypical example is the singleton set consisting of the state $\vec{A}$ in which all players play A in a model where G has (A, A) as a symmetric Nash equilibrium, and where players do not switch away from a best response unless a mutation occurs.

While it is common in the literature to examine only the long run behavior of models, this is not because the medium run is not an important consideration. As Ellison (1993a) and others have noted, a description of the implications of evolution which includes only an identification of the long run equilibria may be misleading in that what happens in the "long run" is often very different from what we would expect to observe in, say, the first

[3]The assumptions on the nature of the perturbations we have made are sufficient to ensure that the limit does in fact exist. To see this note that for any two states x and y Lemma 3.1 of Chapter 6 of Freidlin and Wentzell (1984) expresses the quantity $\frac{\mu_x^*}{\mu_y^*}$ as a ratio of polynomials in the transition probabilities. Because the transition probabilities themselves are asymptotically proportional to powers of ϵ, each of these ratios will have a limit as $\epsilon \to 0$.

[4]Such states are also at times referred to as being stochastically stable.

[5]Recall that $\Omega \subset Z$ is a recurrent class if $\forall w \in \Omega$, $\text{Prob}\{z_{t+1}^0 \in \Omega | z_t^0 = w\} = 1$, and if for all $w, w' \in \Omega$ there exists $s > 0$ such that $\text{Prob}\{z_{t+s}^0 = w' | z_t^0 = w\} > 0$.

trillion periods. Many authors have nonetheless focussed on the long run, no doubt in part because this is what the standard techniques allow one to do most easily. The main theorem of this paper provides a general characterization of the rate at which systems converge, with my hope being that this will induce future researchers to provide more complete analyses. In particular, writing $W(x, Y, \epsilon)$ for the expected wait until a state belonging to the set Y is first reached given that play in the ϵ-perturbed model begins in state x, and supposing Ω is the set of long run equilibria of the model, this paper uses $\max_{x \in Z} W(x, \Omega, \epsilon)$ as a measure of how quickly the system converges to its long run limit. If this maximum wait is small, convergence is fast and we will regard Ω not just as the collection of states which continue to occur frequently in the long run, but also as a good prediction for which behaviors we might expect to see in the medium run. If the maximum wait is large, one should be cautious in drawing conclusions from an analysis of long run equilibria. Note that in practice all these waiting times will tend to infinity as ϵ goes to zero. Whether we regard convergence as being fast or slow in a given model will depend on how quickly the waiting times blow up as ϵ goes to zero.

3 The radius and coradius of basins of attraction

This section defines two new concepts, the radius and coradius of the basin of attraction of a limit set, which are central to this paper's characterization of the behavior of evolutionary models with noise. Essentially, the main theorem of this paper provides a sufficient condition for identifying long run equilibria by formalizing a very simple intuitive argument relating long run behavior and waiting times: *if* a model has a collection of states Ω which is very persistent once it is reached, and which is sufficiently attractive in the sense of being reached relatively quickly after play begins in any other state, then in the long run we will observe that states in Ω occur most of the time. The nontrivial work involved in developing this observation into an useful characterization of behavior consists of developing measures of the persistence of and of the tendency of sets to be reentered which are reasonably easy to compute. The particular approach I take is develop measures which reflect critical aspects of the size and structure of the basins of attraction of a model's limit sets. What I see as surprising is not that a characterization along these lines can be developed, but rather that a fairly simple the characterization of this form turns out to be sufficiently powerful to

identify the long run equilibria of most of the models which have been previously analyzed.

I begin with a measure of persistence. Suppose $(Z, P(0))$ is the base Markov process of the model above, and let Ω be a union of one or more of its limit sets. Write $D(\Omega)$ for the set of initial states from which the Markov process converges to Ω with probability one, *i.e.*

$$D(\Omega) = \{z \in Z | \text{Prob}\{\exists T \text{ s.t. } z_t \in \Omega \ \forall t > T | z_0 = z\} = 1\}.$$

The measure of persistence which appears in the main theorem is simply the number of "mutations" necessary to leave a basin of attraction. Formally, define a path from Ω out of $D(\Omega)$ to be a finite sequence of distinct states $(z_1, z_2 \ldots, z_T)$ with $z_1 \in \Omega$, $z_t \in D(\Omega)$ for all $t < T$, and $z_T \notin D(\Omega)$. Write $S(\Omega, Z - D(\Omega))$ for the set of all such paths. The definition of cost can be extended to paths by setting $c(z_1, z_2, \ldots, z_T) = \sum_{t=1}^{T-1} c(z_t, z_{t+1})$. I define the *radius* of the basin of attraction of Ω, $R(\Omega)$, to be the minimum cost of any path from Ω out of $D(\Omega)$, *i.e.*

$$R(\Omega) = \min_{(z_1,\ldots,z_T) \in S(\Omega, Z - D(\Omega))} c(z_1, z_2, \ldots, z_T).$$

Note that if we define set-to-set cost functions by

$$c(X, Y) = \min_{(z_1,\ldots,z_T) \in S(X,Y)} c(z_1, z_2, \ldots, z_T),$$

then $R(\Omega) = c(\Omega, Z - D(\Omega))$.

Geometrically, one can think of the radius of Ω as the radius of the largest neighborhood of Ω from which the Markov process must converge back to Ω. The larger is the radius of the basin of attraction of a limit set, the longer will it take for the Markov process to escape from it. Note that the calculation of the radius of Ω does not require that one explore the full dynamic system; it typically suffices to examine the dynamics in a neighborhood of Ω. In practice, the least costly path from Ω out of $D(\Omega)$ is often a direct path (w, z_2). In each of the examples presented in this paper, the cost function $c(x, y)$ takes the form $c(x, y) = \min_{\{z | P_{zz}(0) > 0\}} d(z, y)$, with d satisfying the triangle inequality. In this case, a simple method for proving that $R(\Omega) = k$ is to exhibit states $w \in \Omega$, and $z_2 \notin D(\Omega)$ with $c(w, z_2) = k$ and to show both that $\min_{w \in \Omega} c(w, z) < k \Longrightarrow z \in D(\Omega)$, and that $\min_{w \in \Omega} c(w, z) < k \Longrightarrow \min_{w \in \Omega} c(w, z') \leq \min_{w \in \Omega} c(w, z)$ for all z' with $P_{zz'}(0) > 0$. To see that this is sufficient, one shows that the least costly path is a direct path by showing inductively that there exist $w_1, w_2, \ldots, w_{T-1} \in \Omega$ such that $c(w, z_1, z_2, \ldots, z_T) \geq$

$c(w_1, z_2, \ldots, z_T) \geq \cdots \geq c(w_{T-1}, z_T)$.[6]

The second property of a union of limit sets Ω which will play a critical role in our main theorem is the length of time necessary to reach the basin of attraction of Ω from any other state. One simple way to put a (not particularly tight) bound on this waiting time is to count the "number of mutations" which are required. Formally, define a path from x to Ω to be a finite sequence of distinct states $(z_1, z_2, \ldots, z_T)$ with $z_1 = x$, $z_t \notin \Omega$ for all $t < T$, and $z_T \in \Omega$. Write $S(x, \Omega)$ for the set of all such paths. I define the *coradius* of the basin of attraction of Ω, $CR(\Omega)$, by

$$CR(\Omega) = \max_{x \notin \Omega} \min_{(z_1, \ldots, z_T) \in S(x, \Omega)} c(z_1, z_2, \ldots, z_T).$$

It simplifies the calculation of the coradius to keep in mind that the maximum in the formula above is always achieved at a state which belongs to a limit set. Intuitively, the coradius of the basin of attraction of Ω is small when all states are, in a cost-of-paths sense, close to that set.

When the coradius of Ω is small it is easy to see that the basin of attraction of Ω will be entered fairly soon after play starts in any other state. Given any initial state, there is a nonnegligible probability of reaching $D(\Omega)$ within some finite number T of periods. If this does not occur, then the period T state is again an inital condition from which there is a nonnegligible probability of reaching $D(\Omega)$ by period $2T$, and so on. The bound on waiting times provided by the coradius, however, is often not very sharp. In particular, the computation does not reflect the fact that evolution toward Ω may be greatly speeded by the presence of intermediate steady states along the path. Defining a variant of the coradius concept which takes this into account will allow us to achieve a tighter bound on return times, producing more widely applicable characterization of long run equilibria.

---

[6] The inductive step consists simply of noting that for $z_1' \in \arg\min_{\{z | P_{z_1 z}(0) > 0\}} d(z, z_2)$, $w_1 \in \arg\min_{w' \in \Omega} c(w', z_1')$, and $w_1^* \in \arg\min_{\{w' | P_{w_1 w'}(0) > 0\}} d(w', z_1')$,

$$\begin{aligned} c(w, z_1, z_2) &= c(w, z_1) + c(z_1, z_2) = c(w, z_1) + d(z_1', z_2) \\ &\geq c(w, z_1) + d(w_1^*, z_2) - d(w_1^*, z_1') \\ &= c(w, z_1) + d(w_1^*, z_2) - c(w_1, z_1') \\ &\geq d(w_1^*, z_2) \geq c(w_1, z_2). \end{aligned}$$

The variant we will use is based on a modified notion of the cost of a path which allocates a portion of the total cost of getting to Ω to the necessity of leaving the basin of attraction of each limit set the path passes through. Specifically, for $(z_1, z_2, \ldots, z_T)$ a path from x to Ω, let $L_1, L_2, \ldots, L_r \subset \Omega$ be the sequence of limit sets through which the path passes consecutively, with the convention that a limit set can be appear on the list multiple times, but not successively. Define a *modified cost function*, c^*, by

$$c^*(z_1, z_2, \ldots, z_T) = c(z_1, z_2, \ldots, z_T) - \sum_{i=2}^{r-1} R(L_i).$$

The definition of modified cost can be extended to a point-to-set concept by setting

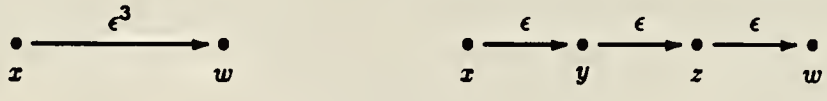
$$c^*(x, \Omega) = \min_{(z_1, \ldots, z_T) \in S(x, \Omega)} c^*(z_1, z_2, \ldots, z_T).$$

The *modified coradius* of the basin of attraction of Ω is then defined by

$$CR^*(\Omega) = \max_{x \notin \Omega} c^*(x, \Omega).$$

Note that $CR^*(\Omega) \leq CR(\Omega)$.

The definition of the modified coradius is meant to capture the effect of intermediate steady states on expected waiting times. To get some intuition for this it is instructive to compare the two simple Markov processes in the diagram below. In the diagram arrows are used to indicate the possible transitions out of each state and their probabilities (with null transitions accounting for the remaining probabilities). In the Markov process on the left, both the cost and the modified cost of the path (x, w) is three, and the expected wait until such a transition occurs is $\frac{1}{\epsilon^3}$. In contrast, while the cost of the path (x, y, z, w) in the Markov process on the right is still three, the expected wait until a transition from x to w occurs is only $\frac{1}{\epsilon} + \frac{1}{\epsilon} + \frac{1}{\epsilon} = \frac{3}{\epsilon}$. The modified cost of the path is also one.



A biological analogy provides some intuition for the result. Think of the graphs as representing two different evironments in which three major genetic mutations are necessary

to produce the more fit animal w from animal x. (For example, to produce a bat from a mouse.) The graph of the left reflects a situation in which an animal with any one or two of these mutations could not survive. In this case, getting the large mutation we need (growing an entire wing?) is a very unlikely event. In the process on the right, each single mutation on its own provides an increase in fitness which allows that mutation to take over the population. Clearly, the large cumulative change from x to w seems relatively more plausible when such gradual change is possible.

From the comparison of the two Markov processes above, it should be clear that intermediate limit sets can speed evolution. Why the particular correction for this embodied in the modified coradius definition is approprate is much less obvious. While Theorem 1 shows formally that the modified coradius can be used to bound return times, an example may be more useful in presenting the basic intuition. In the three state Markov process shown below, suppose that $c(x,y) > r_y$ and $c(y,\Omega) > r_y$ so that Ω is the long run equilibria with $R(\Omega) = \infty$ and $CR^*(\Omega) = c(x,y) + c(y,\Omega) - r_y$. To compute $W(x,\Omega,\epsilon)$ (which will turn out to be the longest expected wait) note that

$$W(x,\Omega,\epsilon) = E(N_x) + E(N_y),$$

with N_x and N_y being the number of times x and y occur before Ω is reached conditional on the system starting in state x. We can write $E(N_x)$ as the product of the expected length of the run of x's before an $x \to y$ transition occurs and the number of times the transition $x \to y$ occurs. Hence,

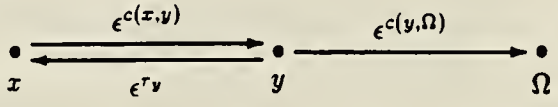
$$\begin{aligned} E(N_x) &= (1+\epsilon^{-c(x,y)})\left(1+\frac{\epsilon^{r_y}+\epsilon^{c(y,\Omega)}}{\epsilon^{c(y,\Omega)}}\right) \\ &\sim \epsilon^{-(c(x,y)+c(y,\Omega)-r_y)}. \end{aligned}$$

Similarly, we can write $E(N_y)$ as the product of the length of each run of y's and the number of $x \to y$ transitions. This gives $E(N_y) \sim \epsilon^{-(r_y+c(y,\Omega)-r_y)}$, which is asymptotically negligible compared to $E(N_x)$. Hence we have

$$W(x,\Omega,\epsilon) \sim \epsilon^{-(c(x,y)+c(y,\Omega)-r_y)},$$

which is exactly $\epsilon^{-CR^*(\Omega)}$. Looking carefully at this calculation, one can see that the subtraction of the radius of the intermediate limit set reflects the fact that because most of

the time the system is in state x the waiting time is essentially determined by the wait to leave x and the probability of a transition from y to Ω *conditional on a transition out of* y *occurring*. While the unconditional probability of the transition is $\epsilon^{c(y,\Omega)}$, the conditional probability is $\epsilon^{c(y,\Omega)-r_y}$.



4 The main theorem

We are now in a position to present the theorem which contains the main general results of this paper. The theorem provides a description of both long run and medium run behavior in evolutionary models with noise (using the radius and modified coradius measures). I hope that this theorem is seen as valuable in providing both a general tool for analyzing the medium run and intuition for why the results of previous papers are what they are. I begin with the theorem and its proof, and then illustrate the use of the theorem and discuss some of its limitations via a number of simple examples.

Theorem 1 *In the model described above if* $R(\Omega) > CR^*(\Omega)$, *then*

(a) All of the long run equilibria of the system are contained in Ω.

(b) For any $y \notin \Omega$, *we have* $W(y,\Omega,\epsilon) = O(\epsilon^{-CR^*(\Omega)})$ *as* $\epsilon \to 0$.

Note that as a corollary, the same results hold if $R(\Omega) > CR(\Omega)$.

Proof

Write $\mu^\epsilon(z)$ for the probability assigned to state z in the steady-state distribution of the ϵ-**perturbed** model. For part (a) it suffices to show that $\mu^\epsilon(y)/\mu^\epsilon(\Omega) \to 0$ as $\epsilon \to 0$ for all $y \notin \Omega$. A standard characterization of the steady-state distribution is that

$$\frac{\mu^\epsilon(y)}{\mu^\epsilon(\Omega)} = \frac{\text{E \# of times } y \text{ occurs before } \Omega \text{ is reached starting at } y}{\text{E \# of times } \Omega \text{ occurs before } y \text{ is reached starting in } \Omega},$$

with the expectation in the denominator being also over starting points in Ω.[7] The numerator of the RHS is at most $W(y,\Omega,\epsilon)$, while the denominator is bounded below as $\epsilon \to 0$ by

[7] See *e.g.* Theorem 6.2.3 of Kemeny and Snell (1960).

a nonvanishing constant times $\min_{w\in\Omega} W(w, Z - D(\Omega), \epsilon)$ (because almost all of the time spent in $D(\Omega)$ is spent in Ω.) Hence to prove both part (a) and part (b) of the theorem it suffices to show that

$$W(w, Z - D(\Omega), \epsilon) \sim \epsilon^{-R(\Omega)} \qquad \forall w \in \Omega,$$

and

$$W(y, \Omega, \epsilon) = O(\epsilon^{-CR^*(\Omega)}) \qquad \forall y \notin \Omega.$$

The first of these asymptotic characterizations of waiting times, that $W(w, Z-D(\Omega), \epsilon) \sim \epsilon^{-R(\Omega)}$ for $w \in \Omega$ is fairly obvious. To see that $W(w, Z - D(\Omega), \epsilon) \leq K\epsilon^{-R(\Omega)}$ note first that we may find constants T and c_1 such that for any $w \in D(\Omega)$ there exists a path $w = z_1, z_2, \ldots, z_T$ with $z_T \in Z - D(\Omega)$ such that the product of all the transition probabilities on the path is at least $c_1\epsilon^{R(\Omega)}$. Conditioning on the outcome of the first T periods we have

$$W(w, Z - D(\Omega), \epsilon) \leq T + (1 - c_1\epsilon^{R(\Omega)})\mathrm{Max}_{z\in D(\Omega)} W(z, Z - D(\Omega), \epsilon)$$

Taking the maximum of both sides over all $w \in D(\Omega)$ yields

$$\mathrm{Max}_{w\in D(\Omega)} W(w, Z - D(\Omega), \epsilon) \leq (T/c_1)\epsilon^{-R(\Omega)}.$$

To see that $W(w, Z - D(\Omega), \epsilon) \geq k\epsilon^{-R(\Omega)}$ for any $w \in \Omega$, note that

$$\begin{aligned} W(w, Z - D(\Omega), \epsilon) &\geq \mathrm{E}(\#\text{ times in } \Omega \text{ before } Z - D(\Omega) \text{ is reached } | z_0 = w) \\ &\geq 1/\mathrm{Max}_{y\in\Omega}\mathrm{Prob}\{Z - D(\Omega) \text{ is reached before returning to } \Omega | z_0 = y\}. \end{aligned}$$

Given that the system returns to Ω in a finite number of periods (in expectation) conditional on less than $R(\Omega)$ mutations occuring, the probability of going from w to $Z - D(\Omega)$ without hitting Ω is clearly of exact order $\epsilon^{R(\Omega)}$, and hence we have the desired lower bound for $W(w, Z - D(\Omega), \epsilon)$ as well.

It remains now only to show that

$$\mathrm{Max}_{y\notin\Omega} W(y, \Omega, \epsilon) = O(\epsilon^{-CR^*(\Omega)}).$$

To begin, we note that if we define set-to-set expected waiting times on a worst case basis,

$$W(X, Y, \epsilon) = \max_{x\in X} W(x, Y, \epsilon),$$

and write $\mathcal{L}$ for the union of limit sets of $(Z, P(0))$, then because in the limit as $\epsilon \to 0$ the expected wait until a limit set is first reached is finite, it will suffice to show that

$$\text{Max}_{y \in \mathcal{L}-\Omega} W(y, \Omega, \epsilon) = O(\epsilon^{-CR^*(\Omega)}).$$

Write $\overline{W}(\Omega, \epsilon)$ for $\text{Max}_{y \in \mathcal{L}-\Omega} W(y, \Omega, \epsilon)$.

For each $y \in \mathcal{L} - \Omega$, let $y = z_1, z_2, \ldots, z_T$ be a path from y to Ω of minimum possible modified cost. Because any element of a limit set can be reached at zero cost from each other element of the same limit set, we may choose this path so that each of the limit sets $L_1, L_2, \ldots, L_r$ through which it passes is distinct, and such that the path is contained in each of these limit sets for a set of successive periods. (Remember that $y \in L_1$). For this path to have minimum modified cost it must be the case that

$$c(z_1, z_2, \ldots, z_T) = \sum_{i=1}^{r-1} c(L_i, L_{i+1}),$$

and if $z_t \in L_i$ it must also be the case that

$$c^*(z_t, z_{t+1}, \ldots, z_T) = c^*(L_i, \Omega).$$

(Otherwise, a lower modified cost path can be constructed by concatenating $(z_1, \ldots, z_t)$ with a minimum modified cost path from L_i to Ω.)

Now, the crucial element of the proof is to simply expand the expected waiting time to reach Ω from L_1 as the expected wait to reach any other limit set, plus the expected wait to reach Ω from that point. Writing $q(z|y)$ for the probability that the first element of $\mathcal{L} - L_1$ reached is z (after starting at y), and assuming that y is the element of L_1 which has the longest expected wait to reach Ω we have

$$W(L_1, \Omega, \epsilon) = W(L_1, \mathcal{L} - L_1, \epsilon) + \sum_{z \in \mathcal{L}-L_1} q(z|y) W(z, \Omega, \epsilon).$$

Writing q_{12} for $\min_{\ell \in L_1} \sum_{z \in L_2} q(z|\ell)$ we have

$$W(L_1, \Omega, \epsilon) \leq W(L_1, \mathcal{L} - L_1, \epsilon) + q_{12} W(L_2, \Omega, \epsilon) + (1 - q_{12}) \overline{W}(\Omega, \epsilon).$$

Repeating this process we get

$$W(y, \Omega, \epsilon) \quad \leq \quad W(L_1, \mathcal{L} - L_1, \epsilon) + (1 - q_{12}) \overline{W}(\Omega, \epsilon)$$

$$
\begin{aligned}
&+q_{12}(W(L_2,\mathcal{L}-L_2,\epsilon)+(1-q_{23})\overline{W}(\Omega,\epsilon))\\
&+q_{12}q_{23}(W(L_3,\mathcal{L}-L_3,\epsilon)+(1-q_{34})\overline{W}(\Omega,\epsilon))\\
&+\cdots\cdots \qquad\qquad \cdots\\
&+q_{12}q_{23}\cdots q_{r-2r-1}(W(L_{r-1},\mathcal{L}-L_{r-1},\epsilon)+(1-q_{r-1r})\overline{W}(\Omega,\epsilon))\\
=\;& W(L_1,\mathcal{L}-L_1,\epsilon)+q_{12}W(L_1,\mathcal{L}-L_2,\epsilon)+q_{12}q_{23}W(L_3,\mathcal{L}-L_3,\epsilon)\\
&+\ldots+(1-q_{12}q_{23}\cdots q_{r-1r})\overline{W}(\Omega,\epsilon)
\end{aligned}
$$

To show that $\overline{W}(\Omega,\epsilon)=O(\epsilon^{-CR^*(\Omega)})$ it suffices to show that this holds on each of the subsets of ϵ values for which $W(y,\Omega,\epsilon)=\overline{W}(\Omega,\epsilon)$. For such ϵ's the expression above gives

$$
\overline{W}(\Omega,\epsilon)\leq\frac{W(L_1,\mathcal{L}-L_1,\epsilon)}{q_{12}q_{23}\cdots q_{r-1r}}+\ldots+\frac{W((L_{r-1},\mathcal{L}-L_{r-1},\epsilon)}{q_{r-1r}}.
$$

To show that $\overline{W}(\Omega,\epsilon)=O(\epsilon^{-CR^*(\Omega)})$ it now suffices to show that each of the expressions on the RHS of the above equation is $O(\epsilon^{-CR^*(\Omega)})$.

To do so, consider the elements which appear in the expressions. Because the minimum cost path from each $z\in D(L_i)$ to L_{i+1} has cost $c(L_i,L_{i+1})$ and $W(L_i,Z-D(L_i),\epsilon)\sim\epsilon^{-R(L_i)}$ we have $q_{ii+1}\sim\epsilon^{c(L_i,L_{i+1})-R(L_i)}$. Because the unperturbed Markov process converges to a limit set in finite time from any point and the probability of converging to L_i is uniformly bounded away from zero for any $y\notin D(L_i)$, we also have $W(L_i,\mathcal{L}-L_i,\epsilon)\sim W(L_i,Z-D(L_i),\epsilon)$.

Putting these together we have

$$
\begin{aligned}
\frac{W(L_i,\mathcal{L}-L_i,\epsilon)}{q_{ii+1}\cdots q_{r-1r}} &= \frac{\epsilon^{-R(L_i)}}{\epsilon^{c(L_i,L_{i+1})-R(L_i)}\cdots\epsilon^{c(L_{r-1},L_r)-R(L_{r-1})}}\\
&= \epsilon^{-(\sum_{j=i}^{r-1}c(L_j,L_{j+1})-\sum_{j=i+1}^{r-1}R(L_j))}\\
&= \epsilon^{-c^*(L_i,\Omega)},
\end{aligned}
$$

with the last line following from our earlier discussion of the nature of minimum modified cost paths. For each i this expression is $O(\epsilon^{-CR^*(\Omega)})$ and hence we have $W(y,\Omega,\epsilon)=O(\epsilon^{-CR^*(\Omega)})$, the second main component of the proof.

QED.

At this point, I would like to give some motivation for the theorem by contrasting the analysis it provides with the now standard techniques developed in KMR, Young (1993a), and Kandori and Rob (1992). These techniques provide an algorithm which in principle will

find the long run equilibria of any model which fits in our framework. A direct application of the algorithm requires one to first identify all of the limit sets of the model and find the tree on the set of limit sets which minimizes a particular cost function.

The obvious drawback of this paper's characterization of long run equilibria in comparison with the standard techniques is that it is *not* universal. However, as is discussed in Section 9, its restricted applicability may not be such a serious limitation from a practical viewpoint in that the set of models to which the main theorem applies appears to contain most models for which economists have previously succeeded in applying the standard techniques. The main theorem of this paper has several potential advantages. First, and most concretely, the theorem provides a convergence rate as well as a long run limit. Second, the theorem provides a clear intuition for why it works, whereas the results of previous papers are often regarded as being mysterious. My guess, in fact, is that what readers will find surprising about the intuition provided by this paper is not that a characterization of long run equilibrium can be derived from it, but rather that there is nothing more going on in previous papers. Finally, the standard techniques are ill-suited to direct application to models which contain a large number of limit sets or when one wants to derive theorems which apply to classes of models broad enough as to include members for which the limit sets differ. It is my hope that the fact that the techniques of this paper may facilitate such analyses will be seen as one of its primary contributions.[8]

The basic plan for the remainder of this section (and for the remainder of the paper) is to present a sequence of examples which are intended in varying proportions to illustrate the use of the main theorem and to present applications which are of interest in their own right. I begin with a simple example which illustrates the use of the radius, coradius and modified coradius measures, and which provides some geometric intuition.

Example 1 ***Suppose N players are repeatedly randomly matched to play the game G as in the model of Section 2 with uniform matching best-reply dynamics and independent random mutations, i.e. with with player i in period t playing a best response to the distribution of***

[8]The sense in which this paper might facilitate such analyses deserves some discussion. The proof of Theorem 2 clearly shows that all models to which Theorem 1 applies not only could have been handled with the standard techniques via "tree surgery" arguments, but that this could have been done in a fairly systematic way. I would interpret this result as indicating that one can alternately think of this paper as contributing to the analyses of such models by systematizing the use of tree surgery arguments.

strategies in the population in period $t-1$ with probability $1-2\epsilon$, and playing each of his other two strategies with probability ϵ. If G is the game shown on the left below, then for N sufficiently large we have $\mu^*(\vec{A}) = 1$. If G is the game shown on the right below, then for N sufficiently large we have $\mu^*(\vec{B}) = 1$.

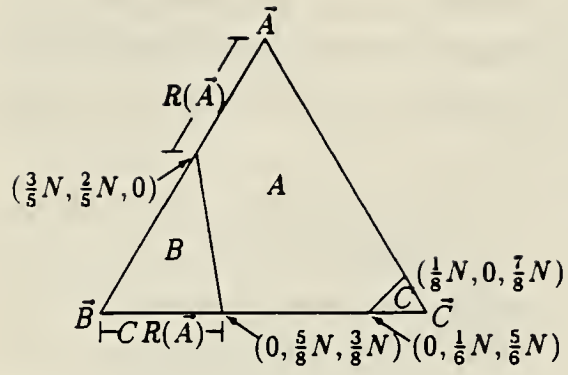
	A	B	C
A	7, 7	5, 5	5, 0
B	5, 5	8, 8	0, 0
C	0, 5	0, 0	6, 6

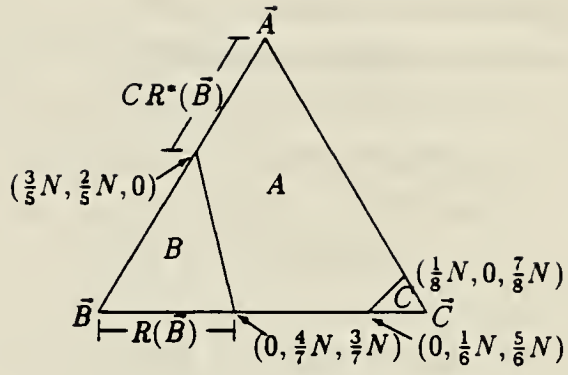
	A	B	C
A	7, 7	5, 5	5, 0
B	5, 5	8, 8	1, 0
C	0, 5	0, 1	6, 6

Proof

The diagrams below show the best-response regions corresponding to each of the games in $(\sigma_A, \sigma_B, \sigma_C)$-space. The deterministic dynamics in each case consist of immediate jumps from each point to the vertex of the triangle corresponding to everyone playing the best response. From the figure on the left, it is fairly obvious that in the first game $R(\vec{A})$, the minimum distance between A and any point not in $D(\vec{A})$ (here $D(\vec{A})$ is approximately the region where A is the best response), is about $\frac{2}{5}N$. The maximum distance between any state and $D(\vec{A})$ (here the distance from $\vec{B}$ to $D(\vec{A})$) is approximately $\frac{3}{8}N$. For this reason, $R(\vec{A}) > CR(\vec{A})$ for N large and the first result follows from Theorem 1. Formalizing this argument requires only a straightforward but tedious accounting for integer problems (and if desired for players not including their own play in the distribution to which they play a best response).

In the figure on the right, the modified coradius measure is useful. Here, $R(\vec{B})$ is about $\frac{3}{7}N$, while $CR^*(\vec{B})$ is about $\frac{2}{5}N$ because this is the modified cost of the path $(\vec{A}, \vec{B})$ and the modified cost of the indirect path $(\vec{C}, \vec{A}, \vec{B})$ is only about $\frac{1}{8}N$. Hence, the result again follows from Theorem 1. Note that in this case a simple radius-coradius argument does not suffice — $CR(\vec{B})$ is approximately $\frac{4}{7}N$.

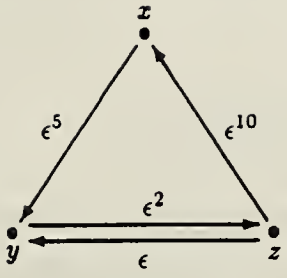


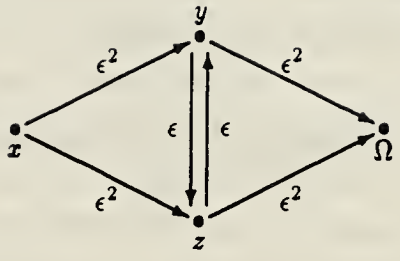


QED.

The bound on the convergence rate in these games is that the expected waits until reaching the long run equilibria are at most of order $\epsilon^{-(3/8)N}$ and $\epsilon^{-(2/5)N}$. There is nothing particularly interesting about this result, other than that as is typical in models with uniform matching the convergence time is rapidly increasing in the population size. The reader should keep in mind that the above examples were chosen largely for their simplicity, and that they are not representative of the situations where the radius-coradius approach of this paper is most useful. The power of the modified coradius measure is much more evident in models with many steady states.

At this time, providing examples which point out a couple of weaknesses of the theorem will probably help the reader to develop a more complete understanding how it works. I would like to remind the reader that unlike the Freidlin-Wentzell characterization, the main theorem of this paper does not provide a characterization which in theory can identify the long run equilibria of all systems. To see this look at the three state Markov process on the left below. It is easy to verify that $R(x) = 5 < CR^*(x) = 11$, $R(y) = 2 < CR^*(y) = 5$, and $R(z) = 1 < CR^*(z) = 5$. The theorem thus does not help us find the long run equilibrium (which turns out to be y). More generally the theorem will never apply unless the long run equilibrium is also the state with the largest radius.





Next, the theorem provides only an upper bound on the expected wait until the system reaches the long run equilibrium, and this bound is not necessarily tight. This point is illustrated by the four-state Markov process on the right above. In this system, Ω is the long run equilibrium and $CR^*(\Omega) = 3$. (Any path from x to Ω has modified cost three.) The expected wait until Ω is reached conditional on starting at x, however, is only $\frac{3}{2\epsilon^2}$, because it takes $\frac{1}{2\epsilon^2}$ periods to reach the set $\{y, z\}$ and then a further $\frac{1}{\epsilon^2}$ periods to reach Ω. Intuitively, the reason why the modified coradius bound is not tight here is that the general calculations which derive this bound do not take full advantage of the relationships between the limit sets. In particular, the calculation is done always assuming that the "worst case" occurs whenever a transition to a new limit set does not correspond to the minimum modified cost path. In this case convergence is faster than such a worst case calculation indicates, because the transitions from y or z which do not reach Ω never take us back to x. A failure to provide tight bounds in some situations seems unavoidable in any computation which does not require a full identification of all of the limit sets and an analysis of the relationships between them.

The example on the left above also provides a nice opportunity to point out a feature of the theorem which has not been emphasized so far — that the Ω in the theorem can be a union of limit sets rather than a single limit set. In this example note that $R(\{x, y\}) = 2 > 1 = CR(\{x, y\})$, and that $R(\{y, z\}) = 10 > 5 = CR(\{y, z\})$. The theorem then implies that all long run equilibria are contained in both $\{x, y\}$ and $\{y, z\}$, so that y is the unique long run equilibrium. While the extent to which being able to make arguments like this is useful is not clear to me, the fact that the theorem applies to unions of limit sets is of practical importance because in models involving extensive form games, *e.g.* Nöldeke and Samuelson (1993, 1994) and Huck and Oechssler (1995), it appears to be common for the

set of long run equilibria to be a union of several limit sets between which the system can move with one or more mutations.

5 Generalizing risk-dominance: $\frac{1}{2}$-dominance

In 2×2 games, the models of KMR and Young (1993a) provide an elegant and robust characterization of the long run impact of mutations on the development of social conventions – societies are led to adopt the risk-dominant equilibrium. This section discusses the generalization of this result, with the primary positive result being that a refinement of risk dominance, $\frac{1}{2}$-dominance, provides a sufficient condition for determining the long run equilibrium.

For the purposes of this section it is necessary to specify a KMR-style model in a bit more detail. Let G be a symmetric $m \times m$ game with S the set of pure strategies. Following KMR suppose that N players are uniformly randomly paired to play G in periods $t = 1, 2, \ldots$. Let $z_t \in Z$ be the period t state which we will assume corresponds simply to an action profile $(s_{1t}, s_{2t}, \ldots, s_{Nt})$. Let P be the Markov transition matrix on Z capturing the reduced form implications of some set of behavioral rules. Let the perturbed process $P(\epsilon)$ be that which results from each player independently following his behavior rule with probability $1 - \epsilon$ and choosing a strategy at random with probability ϵ (with all strategies having positive probability.)

Also following KMR, this section will focus on a particular class of behavior rules. Let σ_{it} be the mixed strategy in which the probability of action s being being played is $\frac{1}{N-1}$ times the number of players $j \neq i$ with $s_{jt} = s$. Let $BR(z_t)$ denote the set of strategies which are a best response to σ_{it} for some i. The (unperturbed) process governing changes in play over time is said to be *Darwinian* if whenever $BR(z_t)$ is a singleton, $P_{z_t z'} > 0$ only if z' has more players playing strategy $BR(z_t)$ than does z_t (or z' has all players playing $BR(z_t)$). The definition is motivated by the supposition that if player i has the opportunity to switch strategies after period t, a reasonable thing for a myopic player to do would be to play a best response to the strategies used by the other players in the most recent period.[9]

Recall that in a symmetric 2×2 game with pure strategy equilibria (A, A) and (B, B),

[9] One example of a Darwinian dynamic is the best-response dynamic under which all players play a best response to the previous period's strategy distribution, *i.e.* $s_{it+1} \in \text{argmax}_s\, g(s, \sigma_{it})$.

the equilibrium (A, A) is said to be *risk-dominant* if A is the best response to the mixed strategy $\frac{1}{2}A + \frac{1}{2}B$. The main result of KMR is that for N sufficiently large the unique long run equilibrium of the above model involves all players playing the risk-dominant equilibrium. Note that this selection is quite robust in that it holds for all Darwinian dynamics, even those which are rigged so that evolution toward one strategy is much faster than evolution in the reverse direction. Ellison (1993) further demonstrates the robustness of this selection in showing that risk-dominant equilibria are selected also in a local interaction model involving k-nearest neighbor matching on a circle and best-reply dynamics.

Given that Harsanyi and Selten's definition of risk dominance appeared in a book titled *A General Theory of Equilibrium Selection in Games*, the obvious first direction for work on generalizations to explore was to see if their definition corresponded to what was selected. Young (1993a) immediately showed that it did not, exhibiting a 3×3 game where in which the risk dominant equilibrium was not selected by his dynamics.[10].

It turns out that not only is Harsanyi and Selten's definition not the sought after generalization, but that no general and robust answer exists. One problem noted by Ellison (1993) is that in 3×3 games the selected equilibrium is dependent on the nature of the matching rule, in particular it may be differ between the uniform and two neighbor matching models. I would like to present now an example illustrating an even more basic nonrobustness — that in 3×3 games the selected equilibrium may differ for different Darwinian dynamics.[11]

Example 2 *For the model described above with G the first 3×3 game discussed in Example 1 the unique long run equlibrium is $\vec{A}$ under the best-reply dynamic, and $\vec{B}$ under the Darwinian dynamic which differs from the best-reply dynamic only when both B and C are being used and all players have A as their best response, in which case only the players playing C are assumed to switch to A.*

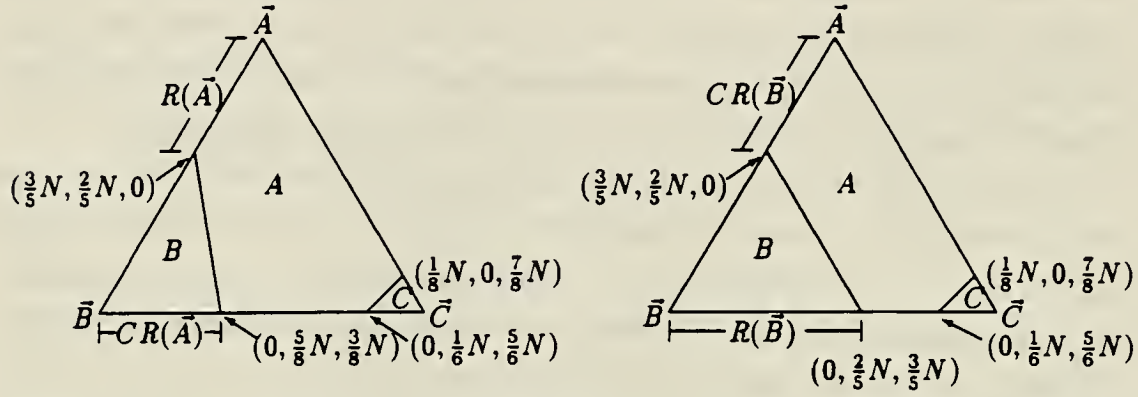
Proof

The results are clear from an examination of the basins of attraction of the equilibria for the two models pictured below. The diagram on the left (repeated from the discussion of Example 1) shows the basins for the best-reply dynamic from which it is easy to see that $R(\vec{A}) > CR(\vec{A})$. The basin of attraction of $\vec{B}$ is larger under the other dynamic, because

[10] The example being the game on the left in Example 1

[11] Hahn (1995) has noted the same result for 2×2 games in a two population model.

starting at some points which were formerly just to the right of the basin of attraction of $\vec{B}$, the unperturbed dynamics now reach $\vec{B}$ in two steps — first moving to a state on the edge where all players play A or B, and then to $\vec{B}$. As a result we have $R(\vec{B}) > CR(\vec{B})$ and $\vec{B}$ is the long run equilibrium.



While robust results can not be obtained generally (or even for the class of games for which risk-dominant equilibria exist), the main result of this section does provide a robust generalization for a substantial class of games. Recall that in $N \times N$ games, Harsanyi and Selten (1988) call the Nash equilibrium (A, A) the risk dominant equilibrium if A is a better response than s to $\frac{1}{2}A + \frac{1}{2}s$ for all pure strategies s such that (s, s) is also a Nash equilibrium. Morris, Rob and Shin (1995) call a symmetric equilibrium (A, A) p-dominant if A is a strict best response against any mixed strategy placing probability at least p on A. Note that $\frac{1}{2}$-dominance is a refinement of risk dominance. The following Corollary contains the main result of this section — that the selection of risk-dominant equilibria in 2×2 games generalizes to the selection of $\frac{1}{2}$-dominant equilibria whenever they exist.

Corollary 1 *In the model above, suppose (A, A) is a $\frac{1}{2}$-dominant equilibrium of G. Then for N sufficiently large the limit distribution μ^* corresponding to any Darwinian dynamic satisfies $\mu^*(\vec{A}) = 1$.*

Proof

The proof is quite easy (with its length here being due to its being presented in extreme detail). Note that since A is a strict best response to any distribution placing probability at least $\frac{1}{2}$ on A there exists a $q < \frac{1}{2}$ such that A is also a strict best response to any distribution placing probability at least q on A. The first step of the proof is to show that

$R(\vec{A}) > (1-q)(N-1)$. To see this, note that whenever $z \neq \vec{A}$ has at least $q(N-1)+1$ players playing A, each player sees at least $q(N-1)$ of the $N-1$ others playing A. Thus, each of them has A as an unique best response and the Darwinian property implies that any state which can be reached at cost zero has strictly more players playing A (*i.e.* we have shown $c(\vec{A}, z') < c(\vec{A}, z)\ \forall z, z'$ with $c(\vec{A}, z) < q(N-1)+1$ and $P_{zz'}(0) > 0$). Iterating the argument above shows that any such z is in $D(\vec{A})$. From the argument following the definition of the radius, these conditions are sufficient to show $R(\vec{A}) > N - (q(N-1)+1)$.

Next, observe that $CR(\vec{A}) \leq q(N-1)+1$, because any state in which at least $q(N-1)+1$ players play A has $\sigma_{it}(A) > q$ for all i, and hence as above is in $D(\vec{A})$. The direct path from any state x to a state in which the first $q(N-1)+1$ players play A while the rest play as in some state which has positive probability in the unperturbed dynamics is a path from x to $D(\vec{A})$ of cost at most $q(N-1)+1$.

Taking N sufficiently large so that $(1-2q)(N-1) \geq 1$, $R(\vec{A}) > CR(\vec{A})$ and Theorem 1 applies.

QED.

I hope that the fact that the proof of the Corollary is fairly trivial will be taken as evidence that the main theorem of the paper can facilitate the development of general results. Because I realize that one might be tempted to conclude instead that the result itself is trivial, I would like to note that both Kandori and Rob's (1992) result on pure coordination games and their (1993) result on risk-dominance in games with the total bandwagon property are immediate corollaries of Corollary 1. In a game where there is a different payoff to coordinating on each of the available actions and the payoffs are zero whenever the players do not coordinate, the pareto optimal equilbrium is also $\frac{1}{2}$-dominant. In a game with the total bandwagon property, risk dominant equilibria are automatically $\frac{1}{2}$-dominant.[12] The result of Kim (1995) on the behavior of the KMR model in a class of symmetric I player, two action games also follows from the argument above (given the appropriate definition of $\frac{1}{2}$-dominance).[13]

[12]Note that what Kandori and Rob (1993) show is that risk dominant equilibria are selected in games satisfying both the total bandwagon property and the monotone share property. Evidently, the latter restriction is unnecessary.

[13]The appropriate definition for the situation being that a symmetric equilibrium is $\frac{1}{2}$-dominant if each player's strategy is his unique best response to any distribution of his opponents' play induced by random

6 Local interaction: two dimensional lattices and fast vs. slow evolution

This section discusses the behavior of models with local interaction. The most notable result is that $\frac{1}{2}$-dominant equilibria are selected and evolution is fast in a model with a two dimensional interaction structure. The result should be of interest both because it clarifies the role of "contagion" dynamics in producing fast evolution, and because it illustrates much more clearly than have any of the previous examples where and how the *modified* coradius computation can be a powerful tool. I hope that the fact that the example occurs late both in the paper and in this section does not deter the reader from looking at the details of how it works.

Evolutionary models with local interaction are intended to capture social situations in which players interact most often with a small stable set of friends, colleagues, or neighbors. Ellison (1993) argued that such models are interesting not only because such relationships exist in the real world, but because it is only in the context of such models that convergence is fast enough to make evolutionary selection arguments plausible.

The first new result presented here is that for k-nearest neighbor local interaction models on a circle the selection of risk-dominant equilibria in 2×2 games again generalizes to the selection of $\frac{1}{2}$-dominant equilibria when they exist (indicating a further robustness of the selection in these games), and that evolution remains relatively fast. I include the result both because of its inherent interest, and because its very quick proof illustrates how the main theorem helps in developing generalizations and in providing convergence rates.[14]

To state this result formally suppose as in Ellison (1993) that players $1, 2, \ldots, N$ are arranged sequentially around a circle. Consider a $2k$ neighbor version of the best-reply dynamic in which player i in period $t+1$ plays a best response to the distribution σ_{it} formed by taking the average of the play in period t by players $i-k, i-k+1, \ldots, i-1, i+1, \ldots, i+k$. Again take the ϵ-perturbed process to be that given by independent random trembles. The result is

Corollary 2 *Suppose* (A, A) *is a* $\frac{1}{2}$*-dominant equilibrium of* G*. For* N *sufficiently large*

matching within a population in which at least half of the players are playing the equilibrium strategy.

[14] Ellison (1993) only provides an ϵ-order approximations to the second largest eigenvalue for the two neighbor model, and does so by a much longer argument.

the limit distribution μ^* corresponding to the $2k$ neighbors on a circle model with best-reply dynamics has $\mu^*(\vec{A}) = 1$, and for all $z \neq \vec{A}$ we have $W(z, \vec{A}, \epsilon) = O(\epsilon^{-(k+1)})$.

Proof

The argument that $R(\vec{A}) > CR(\vec{A})$ follows directly from the proof of Theorem 1 of Ellison (1993). Any state in which $k+1$ adjacent players play A lies in $D(\vec{A})$ because the cluster of players playing A will grow contagiously. Hence $CR(\vec{A}) \leq k+1$.

For $N \geq (k+1)(k+2)$ a path from $\vec{A}$ out of $D(\vec{A})$ requires a mutation among players $1, 2, \ldots, k+1$, a second mutation among players $k+2, \ldots, 2(k+1)$, etc., and a $k+2^{nd}$ among players $(k+1)^2+1, \ldots, (k+1)(k+2)$. Hence, $R(\vec{A}) \geq k+2$ and again Theorem 1 applies.

QED.

Note that an analysis which is tailored more to a given model's dynamics may provide tighter bounds on convergence times. For example, in a model of eight neighbor matching involving a 2×2 game in which A is the best response whenever three of a player's eight neighbors play A Theorem 1 can be used to show that the expected wait to reach A is not asymptotically ϵ^{-5}, but rather ϵ^{-3}.

My primary motivation for discussing the model above and the one that follows is a desire to comment on what it is makes evolution fast in local interaction models. As noted above, the dynamics of the k-nearest neighbor on a circle model feature the rapid contagious expansion of relatively small clusters of players playing the $\frac{1}{2}$-dominant equilibrium. Such dynamics are present, however, only because of the extreme overlap between a player's neighbors and his neighbors' neighbors in this model, and one might wonder whether fast evolution would go away if we eliminated this unreasonable aspect of the social interaction. The remainder of this section is concerned with one model which eliminates this assumption of extreme overlap — a model in which the players are situated at the vertices of a two-dimensional lattice and interact with their four nearest neighbors.[15]

The best-reply dynamics of such a model differ in interesting ways from those of one dimensional models. Note first that unlike one-dimensional models there is no strong force

[15]See Blume (1993) for a discussion of a two-dimensional Ising model which has a somewhat analagous conclusion about long run behavior, and Anderlini and Ianni (1993) and Blume (1994) for analyses of nonergodic models with a two-dimensional interaction structure.

which ensures homogeneneity of actions here. Instead there may be many steady states in which different actions are taken by different segments of the population. For example, in the standard 2×2 coordination game where the players receive a payoff of two from coordinating on A, one from coordinating on B and zero from miscoordinating, one steady state is for a cluster of four adjacent players (in a "square" configuration) to play A while the rest of the population plays B. Players in the cluster are satisfied with A because two of their four neighbors belong to the cluster, while each player outside the cluster is happy playing B because at least three of his four neighbors is also outside the cluster. Other steady states are obtained, for example, by having variously sized separated blocks of players playing A, or by having vertical or horizontal "stripes" where A is played while the remainder of the population plays B.

The fact that small clusters do not grow contagiously in the model without noise makes for an interesting pattern of evolution in the model with noise. Rather than seeing contagious growth triggered by a few mutations we will instead observe a proces of evolution by agglomeration in which small clusters of players playing the $\frac{1}{2}$-dominant equilibrium will arise from mutations, subsequently grow (and sometimes shrink) slowly as mutations occur at the edge of the cluster, until eventually adjacent clusters start to grow together and take over the population. The Corollary shows that this pattern of growth is sufficiently strong to make evolution fast — evidently the fast evolution of local interaction models is not so much about the contagious spread of strategies as it is about the ability of strategies to gain footholds in small areas.

It is because this model has so many steady states (and cycles) that I feel that it is also the ideal setting in which to illustrate how the main theorem of this paper can be applied. The fact that this model has a large number of steady states (and cycles) would make it very difficult to compute the long run equilibrium by directly constructing the minimum order Freidlin-Wentzell tree (a construction whose first step would be the enumeration of all of the limit sets.) At the same time, the presence of a large number of limit sets is precisely what makes the modified coradius calculation useful. Note that in most of the examples we have seen so far a radius-coradius theorem would have sufficed, and hence one can not see why the modified coradius is important. Here, both the identification of the long run equilibrium and the demonstration that evolution is fast will more thoroughly exploit the

fact that the intermediate limit sets exist and a step-by-step path through these limit sets provides a plausible way for change to occur.

Let me now give a brief formal specification of the model, which will be followed by a somewhat lengthy though not too complicated proof that $\frac{1}{2}$-dominance is again sufficient to determine the long run equilibrium, and that evolution is relatively fast. Suppose that N_1N_2 players are located at the vertices of an $N_1 \times N_2$ lattice on the surface of a torus. A state z of the system is a function $z : \{1,\ldots,N_1\} \times \{1,\ldots,N_2\} \longrightarrow S$ with $z(i,j)$ giving the action taken by the player at location $\{i,j\}$. The deterministic best-reply dynamics with nearest neighbor matching are obtained by assuming that in each period each player plays a best response to the strategies used by his *four* immediate neighbors in the previous period, *i.e.* $z_{t+1} = b(z_t)$ satisfies $z_{t+1}(i,j) \in \text{Argmax}_{s \in S} g(s, \sigma_{ijt})$ where σ_{ijt} is the distribution putting probability $\frac{1}{4}$ on each of $z_t(i-1,j)$, $z_t(i+1,j)$, $z_t(i,j-1)$, and $z_t(i,j+1)$, and we assume again that the perturbed process incorporates independent ϵ-probability mutations.

Corollary 3 *Suppose (A,A) is the $\frac{1}{2}$-dominant equilibrium of G. For $N_1 > 3$ and $N_2 > 3$, the limit distribution μ^* corresponding to the best-reply dynamics of the nearest neighbor matching model on an $N_1 \times N_2$ torus has $\mu^*(\vec{A}) = 1$, and for any $z \neq \vec{A}$ we have $W(z,\vec{A},\epsilon) = O(\epsilon^{-3})$ as $\epsilon \to 0$.*

Proof

First, I show that $R(\vec{A}) \geq \min(N_1,N_2)$. To see this, it is easiest to define a distance measure which provides a lower bound on the number of mutations needed to escape $D(\vec{A})$ from any point. Let $g(z)$ be the smaller of the number of rows in which all players play A and the number of columns in which all players play A. If a given row (column) has all players playing A in z, then all players in that row (column) have at least two neighbors playing A, and hence all of them play A in $b(z)$ as well. From this, we know that $g(b(z)) \geq g(z)$. Further, as each "mutation" can only break up one row and one column, we know that $g(z') \geq g(z) - c(z,z')$. Applying this relationship repeatedly, it follows that if $c(\vec{A},z_2,\ldots,z_T) < \min(N_1,N_2)$ then $g(z_T) \geq 1$. To establish that there is no path from $\vec{A}$ to $Z - D(\vec{A})$ with cost less than $\min(N_1,N_2)$, it thus suffices to show that $g(z) \geq 1 \Longrightarrow z \in D(\vec{A})$.

That $g(z) \geq 1 \Longrightarrow z \in D(\vec{A})$ follows from a straightforward explicit calculation. The intuition for this result is that if all players in a "cross" pattern play A, the set of players

playing A will expand out from the center of the cross until it encompasses the entire population. To see this formally, suppose $g(z) \geq 1$ and assume without loss of generality that $z(i,1) = z(1,j) = A$ for all $i \in \{1,\ldots,N_1\}$ and all $j \in \{1,\ldots,N_2\}$. If in addition $z(i,j) = A$ for all (i,j) with $i+j \leq k$, it is easy to see that $b(z)$ also has the same "cross" of players playing A and has $b(z)(i,j) = A$ for all i,j with $i+j \leq k+1$. The conclusion that $g(z) \geq 1 \Longrightarrow z \in D(\vec{A})$ then follows by induction.

To complete the proof, I show next that $CR^*(\vec{A}) \leq 3$. To do so, I explicitly construct for each $x \notin D(\vec{A})$ a path from x to $D(\vec{A})$. Let $z_1 = x$. Let k_1 be the minimum index such that $b^{k_1}(z_1)$ is a member of a stable limit set. Let $z_t = b^{t-1}(z_1)$ for $t = 2,\ldots,k_1+1$. Each transition so far has cost zero. Let z_{2+k_1} be a state in which players $(1,1)$ and $(2,2)$ both play A, and in which the rest of the population plays as in the state $b(z_{1+k_1})$. Note that $c(z_{1+k_1}, z_{2+k_1}) \leq 2$. So that the path being constructed will have a modified cost of three or less, I will define the path with at most one other transition which is not either a best response or a transiton which escapes a distinct limit set whose cost is the radius of that limit set. It turns out that this is indeed possible, because successively larger stable clusters of players playing A can be formed by adding single mutations to the edge of the cluster.

Formally, let k_2 be the minimum index such that $b^{k_2}(z_{2+k_1})$ is a member of a limit set. Note that as the players $(1,2)$ and $(2,1)$ each have two neigbors playing A in each odd period, and players $(1,1)$ and $(2,2)$ each have two neighbors playing A in each even period along the way, one of these pairs plays A in $b^{k_2}(z_{2+k_1})$ as well. Let $z_{2+k_1+i} = b^i(z_{2+k_1})$ for $i = 1,2,\ldots,k_2$. Note that each of the added transitions will have cost zero.

Let $z_{2+k_1+k_2+1}$ be defined so that all of players $(1,1)$, $(1,2)$, $(2,1)$, and $(2,2)$ play A with the remainder of the players playing as in $b(z_{2+k_1+k_2})$. This transition has modified cost at most one. Repeating the process of adding successive best responses we obtain a state $z_{2+k_1+k_2+k_3}$ belonging to a limit set in which all four of those players play A. (This sequence of transitions described in this paragraph should be skipped with k_3 set to zero if the four players already play A in $z_{2+k_1+k_2}$.)

The state $z_{2+k_1+k_2+k_3}$ must either have all players in rows 1 and 2 playing A, or have players $(1,1),\ldots,(1,a)$ and $(2,1),\ldots,(2,a)$ playing A for some $a \geq 2$, with one of $z_{2+k_1+k_2+k_3}(1,a+1)$ and $z_{2+k_1+k_2+k_3}(2,a+1)$ different from $\vec{A}$. Let $z_{2+k_1+k_2+k_3+1}$ be

defined by adding a single mutation to $b(z_{2+k_1+k_2+k_3})$ so that an extra one of the two players mentioned above plays A. Let $z_{2+k_1+k_2+k_3+2}, \cdots, z_{2+k_1+k_2+k_3+k_4}$ be a sequence of best responses such that $z_{2+k_1+k_2+k_3+k_4}$ is again a member of a limit set. This limit set will have at least one of players $(1, a+1)$ and $(2, a+1)$ playing A if $z_{2+k_1+k_2+k_3}$ had neither playing A, and will have both playing A if $z_{2+k_1+k_2+k_3}$ had one of them playing A. Repeating this process, we eventually obtain a path of modified cost at most three to a limit set in which all players in the first two rows play A. Following the same process for the first two columns we obtain a path of modified cost at most three to a state in which all players in the first two columns also play A. By the result above on "crosses", such a state is in the basin of attraction of $\vec{A}$. Hence, $CR^*(\vec{A}) \leq 3$, which completes the proof.

QED.

While this example does a nice job of illustrating the power of the modified coradius and is I think representative of the situations where it is most useful, the reader should keep in mind also that the modified coradius is needed in order to allow the main theorem to be applied to many of the less complicated models which have appeared in the literature. Some of these cases will be pointed out in Section 8.

7 Another example: cycles as long run equilibria

The example of this section concerns one of the the most interesting aspects of stochastic learning models, their ability to provide a rationale for selecting between multiple strict equilibria. Such an interpretation has been encouraged by Young's (1993a) proof that in weakly acyclic games all long run equilibria are Nash equilibria and also by the reasonableness of the selection criterion in a variety of settings.[16] In light of this, the following example is of interest as a simple illustration of the fact that beyond weakly acyclic games, the term **"long run equilibrium"** may be something of a misnomer in that the selected outcome may very well be a cycle which does not involve the unique Nash equilibrium. In this example, (A, A) is the unique Nash equilibrium, while the long run equilibrium is a cycle in which the players synchronously alternate between B and C.

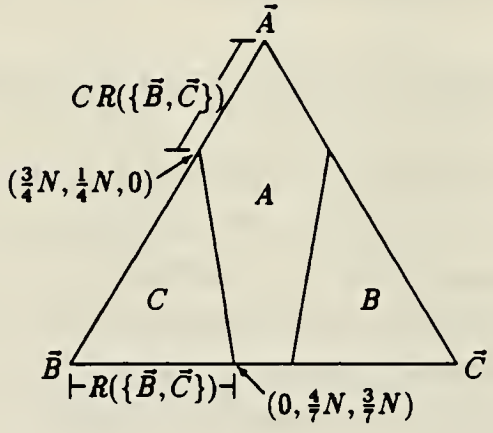
[16] See, for example, Nöldeke and Samuelson (1993, 1994) and Young (1993b)

Example 3 *Suppose N players are repeatedly randomly matched to play the game G shown below as in the model of section 2 with the uniform matching best-reply dynamics. Then for N sufficiently large* $\mu^*(\vec{B}) = \mu^*(\vec{C}) = \frac{1}{2}$.

	A	B	C
A	1, 1	0, 0	0, 0
B	0, 0	-4, -4	3, 3
C	0, 0	3, 3	-4, -4

Proof

Again, the answer is obvious given the structure of the best response regions in $(\sigma_A, \sigma_B, \sigma_C)$-space pictured below. The radius $R(\{\vec{B},\vec{C}\})$ is approximately $\frac{3}{7}N$. The coradius $CR(\{\vec{B},\vec{C}\})$ is approximately $\frac{1}{4}N$. For N large Theorem 1 gives $\mu^*(\vec{B}) + \mu^*(\vec{C}) = 1$, and the result then follows by symmetry.



QED.

Looking at the dynamics above, it is tempting to view the result merely as a curiosity attributable to an unreasonable specification of the deterministic dynamics. If we had instead chosen an unperturbed dynamic in which play changed more gradually or where players had a longer memory, for example, then the cycle would not be a limit set. While the simple example I've given may not be robust, however, the problem is fundamental.

What is selected by models with noise is determined simply by the sizes of and relations between the basins of attraction, and there is no reason to think that being an equilibrium is a necessary condition for having a large basin of attraction. Given that cycles appear as limits not only of the best-reply dynamic, but also of learning processes such as fictitious play, it is not hard to write down a variety of models for which the "long run equilibrium" is a cycle.

8 Relationships with the literature

One of the primary motivations for this paper was a desire to develop a framework which would provide a more intuitive and thorough understanding of the behavior of evolutionary models with noise. To assist the reader in assimilating the ideas of this paper, it is useful to discuss in some detail the relationship between this paper and the existing literature. This section focusses on two topics: the extent to which the main theorem of this paper is applicable to models which have been previously studied and how the techniques of this paper are derived from and build on the previous work of several authors.

Obviously, the framework of this paper is only useful to the extent that it is widely applicable. The question of where the framework is applicable, however, is not as cut and dried as it might seem at first. Depending on why we care about applicability, it may make sense to interpret this question in two very different ways: asking for which models the long run equilibria could have been identified using the main theorem of this paper, or alternatively asking where the theorem further makes the identification of long run equilibria easier (or more general) than it would have been with the standard techniques. I will discuss answers to both forms of the question in this section.

I begin by addressing the first form of the question, with the motivation being that in such cases the results of this paper both provide intuition for why the long run equlibrium is what it is, and allow for a more complete characterization of behavior by providing a bound on the rate of convergence. With this first interpretation of "applicable," one thing we can say is that the theorem of this paper is applicable to almost every model for which long run equilibria have previously been identified.[17] For example, this includes all

[17] One should keep in mind here that this definition of applicability allows me to include many models where the easiest way to see that the the main theorem of this paper applies involves essentially constructing the

of the cases where long run equilibria are identified in Kandori, Mailath, and Rob (1993), Ellison (1993), Evans (1993), Nöldeke and Samuelson (1993, 1994), Robson and Vega-Redondo (1994), Samuelson (1994) and Young (1993a, b).[18] Cases where the theorem is not applicable include both models which do not fit into the framework of Section 2 and models which do fit but for which the theorem simply lacks power. Examples of the first type include models where the noise is sufficiently restricted so as to render the model non-ergodic (as in Anderlini and Ianni (1993) and some of Nöldeke and Samuelson's (1994) signalling games), and papers which employ variants of the solution concept (as in Binmore and Samuelson (1993) who look at an $N \to \infty$ limit). While it is easy to construct simple examples of Markov processes for which the radius-modified coradius calculation simply lacks power (as was done at the end of Section 4), such examples appear to be quite rare in the literature — the only examples I know of are (some of the) differentiated price oligopoly games discussed in Kandori and Rob (1992), for which the tree analysis exploits details of the relationships between the limit sets which are not captured by the modified coradius calculation.

It is my hope that the techniques of this paper will not only allow for a more complete understanding of the previous literature, but will also allow the literature to grow by facilitating the analysis of interesting models which economists have to date shied away from because of their complexity. For this reason, it is useful also to discuss "applicability" in terms of where the techniques of this paper make the analysis easier. The cleanest examples I've found of this type are, not coincidentally, those that I have discussed in this paper. The result on $\frac{1}{2}$-dominance with its trivial proof generalizes several previous results including KMR's original theorem on 2×2 games, Kandori and Rob's (1992) result that pareto optimal equilibria are selected in pure coordination games, and Kandori and Rob's (1993) conclusion that risk dominance is a sufficient condition for being a long run equilibrium in games satisfying their total bandwagon and monotone share properties.[19] Similarly, the

minimum order Freidlin-Wentzell tree and reading the radius and modified coradius off of the tree diagram.

[18]Most of these papers, in fact, require only a $R > CR$ theorem. Young (1993b) and the 3×3 game in Young (1993a) are examples where the modified coradius is necessarily and can be easily determined from the minimum cost tree. The results of the Samuelson and Noldeke-Samuelson analyses are most easily rederived by showing that their lemmas follow from a modified coradius computation.

[19]The result also generalizes the result of Maruta (1995) (which was developed without knowledge of this paper) that $\frac{1}{2}$-dominant equilibria are selected in coordination games.

result on one dimensional local interaction models generalizes the analysis of 2×2 games in Ellison (1993).[20] Generally, it is in allowing one to derive results for broader classes of games (such as games with $\frac{1}{2}$-dominant equilibria as opposed to pure coordination games) and in the analysis of models with many steady states that I would expect the techniques of this paper to prove most useful.

In trying to understand the approach to analyzing models with noise developed here, it is useful also to note how this paper builds on and combines ideas and techniques which have been developed in previous papers. The single most basic idea behind the main theorem of this paper — that the long run equilibrium concept is closely related to waiting times necessary for transitions between limit sets — is new as a basis for an algorithm, but has been clearly recognized as an intuitive description of the long run equilibrium concept from the very beginning (see *e.g.* KMR).

Rather than discussing waiting times, the previous literature on identifying long run equilibria has followed instead the basic approach of analyzing minimum order trees (either directly or via "tree surgery" arguments a la Young (1993a)). In trying to understand how the argument of this paper works, it is useful to note (as is discussed in the section which follows) that the long run equilibrium part of the main theorem of this paper can also be derived fairly easily via a "tree surgery" argument. This argument can be seen as combining two distinct ideas: that evolution with noise tends to favor limit sets with larger basins of attraction, and that the presence of intermediate steady states may speed evolution.

We can find each of these ideas in some form in the previous literature. That a set with a large enough basin of attraction is selected has been previously noted in a number of places. An explicit statement of the fact that $R(\Omega) > CR(\Omega)$ being a sufficient condition for long run equilibrium was independently given by Evans (1993) (see Lemma 3.1), with a tree surgery argument which establishes the result (though it is not stated explicitly) being contained in the proof of Theorem 1 of Ellison (1993). Other authors have noted the limit set with the largest basin of attaction is the long run equilibrium in particular contexts, *e.g.* Kandori and Rob (1993) note that this is the case in 3×3 games satisfying their total bandwagon and monotone share properties. The second idea, that the presence of intermediate steady states may speed evolution, has not previously been stated so explicitly,

[20] Here it is instructive also to compare the tree surgery proof in Ellison (1993) with the fully constructive proof which was given in the working paper version of that paper only for 2-neighbor interaction.

but does clearly play a prominent role in the work of Samuelson (1994) and Nöldeke and Samuelson (1993, 1994). In those papers, the long run equilibria are identified using a lemma which (in my words) states that a component Ω of limit sets receives probability one in the limit distribution if $R(\Omega) > 2$ and Ω can be reached from any other limit set via a chain of single mutations (a condition which ensures that $CR^*(\Omega) = 1$.) The analysis of this paper can be seen as building on this work by noting that a path more generally deserves "credit" for $R(x)$ mutations (rather than one) when passing through the limit set x, and in combining this idea with the previous one so that it can be applied in models where the nonselected limit sets are not so unstable as to be upset by just a single mutation.

The final departure of this paper from the literature is its provision of a general description of medium run behavior in the form of a characterization of waiting times. While no previous papers have provided any general discussions of medium run behavior, the topic has been discussed in a few particular models. Ellison (1993) argues that convergence rates are an important consideration and discusses convergence rates both via an eigenvalue computation and in terms of $N \to \infty$ asymptotics of waiting times. The two subsequent papers which explicitly discuss convergence rates, those of Binmore and Samuelson (1993) and Robson and Vega-Redondo (1994) both take the approach of using minimum T-period transition probabilities to bound waiting times. Though neither paper states a theorem at a level of generality which is greater than that necessary to apply to the model in question, the arguments they give are easily generalized to establish that the waiting time to reach a long run equilibrium Ω which satisfies $R(\Omega) > CR(\Omega)$ is at most $O(\epsilon^{-CR(\Omega)})$. One can think of the recursive waiting time computation of this paper as achieving a tighter bound by taking into account the effects of intermediate steady states.

9 A Freidlin-Wentzell 'tree surgery' argument

In this paper, I have argued that the techniques developed here have several advantages over the Freidlin-Wentzell approach. That the main theorem of this paper provides new intuition and that it is valuable to know how quickly long run equilibria are reached should not be controversial. In addition I have argued that the main theorem of this paper facilitates finding long run equilibria in complex or incompletely specified models. While I hope that the examples presented here provide convincing evidence of this, one could debate this

point by saying that the Freidlin-Wentzell approach is completely general and thus could in principal be applied to any of the models in this paper.

What I would like to note here is that in fact even more is true — the long run equilbrium part of the main theorem can in its full generality be derived as an only somewhat involved application of "tree-surgery" arguments (to borrow a phase from Young (1993a)) to the Freidlin-Wentzell methodology. The primary motivation for doing so is the hope that presenting this connection will improve the understanding of both methodologies.[21] In addition, doing so allows a slight extension of the long run equilibrium part of the theorem to provide partial characterizations and I hope is suggestive of how similar tree surgery arguments may be crafted to solve problems where the main theorem of this paper fails to apply

Theorem 2 *In the model described in Section 2, if for some limit set* Ω *and some state* $x \notin \Omega$ *we have* $R(\Omega) > c^*(x,\Omega)$ *then* $\mu^*(x) = 0$. *If* $R(\Omega) = c^*(x,\Omega)$ *then* $\mu^*(x) > 0$ *implies* $\mu^*(\Omega) > 0$.

Note that part (a) of Theorem 1 follows from the first conclusion of this theorem because $c^*(z,\Omega) \leq CR^*(\Omega)\ \forall z \notin \Omega$, and hence the first conclusion implies $\mu^*(z) = 0\ \forall z \notin \Omega$, which in combination with the fact that $\mu^*(Z) = 1$ implies that $\mu^*(\Omega) = 1$.

Proof

The argument given here follows Foster and Young (1990), Kandori, Mailath, and Rob (1993), and Young (1993a) in relying on Freidlin and Wentzell's (1984) tree characterization of μ^*. The argument differs somewhat in that it is based not as much on an explicit construction of the minimum order trees involved as on a "tree surgery" argument which recognizes that for the purposes of this theorem we care only about whether the minimum order tree has a certain property, and this can be examined by considering modifications of trees without the property. The argument generalizes the proof of Theorem 1 in Ellison (1993) and the main Lemma of Evans (1993).

An x-tree t is a function $t : Z \longrightarrow Z$ such that $t(x) = x$ and such that $\forall z \neq x$ there exists k with $t^k(z) = x$. Define the cost of an x-tree t, $c(t)$, by $c(t) = \sum_{z \neq x} c(z, t(z))$.

[21] In particular, one should note how the arguments here can be seen as combining the surgery technique inherent in Ellison (1993) and developed generally in Evans (1993) with an extension of the single transition techniqne used by Samuelson (1993) and Samuelson and Nöldeke (1993, 1994).

A consequence of Freidlin and Wentzell's results is that long run equilibria correspond to minimum order trees. Specifically, x is *not* a long run equilbrium if for some $w \neq x$ there exists an w-tree t such that all x-trees t' have $c(t') > c(t)$.

To show that $x \notin \Omega$ is not a long run equilibrium when $R(\Omega) > c^*(x, \Omega)$, I show how, given an x-tree t' one can (for some $w \in \Omega$) construct an w-tree t'' which has a strictly lower cost. Suppose t' is an x-tree. Let $(z_1(= x), z_2, \ldots, z_T(= w))$ be a path from x to Ω with minimum modified cost. This cost is at most $c^*(x, \Omega)$. Let $(L_1, L_2, \ldots, L_r = \Omega)$ be the set of limit sets crossed along this path. With an appropriate choice of path we may assume that these limits sets are distinct. Note that for any limit set Ω' of $(Z, P(0))$ and for any state $w' \in \Omega'$ we may define a w'-tree $b : D(\Omega') \to D(\Omega')$ such that $c(z, b(z)) = 0 \; \forall z \neq w'$. Let $Y = \bigcup_{i=1}^{r} D(L_i)$. Combining several maps like that above we may choose $b : Y \to Y$ such that for some $k > 0$ we have $b^k(z) \in \{z_1, z_2, \ldots, z_T\}$ for all $z \in Y$, such that $c(z, b(z)) = 0$ if $z \notin \{z_1, z_2, \ldots, z_T\}$, and such that $b(w) = w$.

Define an w-tree t'' by

$$t''(z) = \begin{cases} z_{t+1} & \text{if } z = z_t \text{ for some } t \in \{1, 2, \ldots, T-1\} \\ b(z) & \text{if } z \in Y, z \notin \{z_1, z_2, \ldots, z_{T-1}\} \\ t'(z) & \text{otherwise} \end{cases}$$

To see that this is a w-tree note that for any state z the path $(z, t''(z), t''^2(z), \ldots)$ coincides with that of t' until it reaches either an element of Y or an element of $\{z_1, z_2, \ldots, z_T\}$. Because t' is an x-tree (and $z_1 = x$), this ensures that one of these sets will be reached. If the second set is reached, the path clearly leads to w. If the first set is reached then, by the definition of b, the second set will be reached within k periods and again the path leads to w.

Because the trees t' and t'' are identical at all states covered by the "otherwise" line of the definition of t'', we may cancel the costs of transition for all such states when comparing the costs of t' and t''. In addition, we know $c(z, b(z)) = 0 \; \forall z \in Y$ with $z \notin \{z_1, z_2, \ldots, z_T\}$. Using these facts we can write

$$c(t'') - c(t') = \sum_{t=1}^{T-1} c(z_t, z_{t+1}) - \sum_{z \in Y} c(z, t'(z)).$$

Now, consider each of the two sums on the right hand side of this expression. The first is at most $c^*(x, \Omega) + \sum_{i=2}^{r-1} R(L_i)$. To bound the second, note that, because t' is an x-tree,

corresponding to any limit set Ω' with $x \notin D(\Omega')$ there exists a $k > 0$ and a $w' \in \Omega'$ such that $(w', t'(w'), \ldots, t'^k(w'))$ is a path from Ω' out of $D(\Omega')$. The cost of such a path is at least $R(\Omega')$. If $\Omega' \in \{L_2, \ldots, L_r = \Omega\}$, the cost of each of the transitions in such a path appears as a term in the second sum on the right hand side. Further, the terms are distinct for distinct limit sets. Hence, that sum is at least $R(\Omega) + \sum_{i=2}^{r-1} R(L_i)$. Hence, we know

$$c(t'') - c(t') \leq c^*(x, \Omega) + \sum_{i=2}^{r-1} R(L_i) - (R(\Omega) + \sum_{i=2}^{r-1} R(L_i)) < 0.$$

Thus, t'' is a lower cost w-tree as desired.

As for the second conclusion of the theorem, note simply that if $c^*(x, \Omega) = R(\Omega)$, then the argument above shows that the minimum cost x-tree is no cheaper than than a w-tree for some $w \in \Omega$.

QED.

10 Conclusion

In this paper, I have discussed the behavior of stochastic models both in general and in several particular examples. With regard to the former, the paper outlines an approach which involves describing the basins of attraction with two new measures, the radius and coradius. The main theorem is applicable to many of the models which have been previously studied and expands our understanding of these models in two ways: it provides a clear intuitive argument which may eliminate much of the mystery left in the wake of Freidlin-Wentzell tree constructions and provides a measure of the rate at which a model converges. The approach is tractable in the examples discussed here despite the fact that the games involved may have best response cycles, and that in one case the dynamics are so complex as to render even a listing of the stable limit sets difficult. As the literature on stochastic evolution continues to grow, economic interest continues to draw researchers to more complex games. Among the recent notable examples are Young's (1993b) analysis of bargaining, Nöldeke and Samuelson's (1994) analysis of signaling games, and Binmore and Samuelson's (1993) "muddling" model. I hope that both the argument of this paper will spur further research into such topics in the future both by reducing the burden of carrying out proofs and by making analyses more complete and more transparent.

This paper is also a paper about models of evolution with noise. In this area, the primary result of the paper is that the selection of risk dominant equilibria in 2×2 games generalizes to the selection of $\frac{1}{2}$-dominant equilibria whenever they exist. In these cases, the long run equilibrium is more robust to the specification of the matching process and the dynamics than can be generally assured. Several other remarks are also worth restating. In games which are not weakly acyclic, long run equilibria need not be Nash equilibria, even when the Nash equilibrium is unique. Despite lacking contagion dynamics, models with two dimensional local interaction can be like one dimensional local models both in their long run limits and in that convergence is very rapid.

The paper may be of more general interest for the insight it provides into the circumstances in which evolutionary change is likely to rapid. The argument that shifts from one equilibrium to another are most likely to be observed in systems which are amenable to gradual change may be applicable in a wide variety of economic contexts — both where gradual change takes the form of shifts which occur first in small subsets of the population and where it involves a continuous variable adjusting slowly between two extremes.

In the future, the results of this paper might be extended in a number of directions. Among the most promising topics to explore are the possibility of extending the applicability of the theorem by grouping limit sets, the possibility of developing tighter bounds on waiting times by taking advantage some specific features of the dynamics rather than always assuming the worst case, and the use of tree surgery arguments in situations where the main theorem of this paper does not apply.

References

Anderlini, L. and Ianni, A. (1993): "Path Dependence and Learning from Neighbors," mimeo, Cambridge University.

Bergin, J., and Lipman, B. (1993): "Evolution with State-Dependent Mutations," mimeo, Queens University.

Binmore, K., and Samuelson, L. (1993): "Muddling Through: Noisy Equilibrium Selection," mimeo, University of Wisconsin.

Blume, L. (1993): "The Statistical Mechanics of Strategic Interaction," *Games and Economic Behavior* 5, 387-424.

Blume, L. (1994): "The Statistical Mechanics of Best-Response Strategy Revision," mimeo, Cornell University.

Canning, D. (1992): "Average Behavior in Learning Models," *Journal of Economic Theory* 57, 442-472.

Ellison, G. (1993): "Learning, Local Interaction, and Coordination," *Econometrica* 61, 1047-1071.

Evans, R. (1993): "Observability, Imitation and Cooperation in the Repeated Prisoners' Dilemma", working paper, University of Cambridge.

Foster, D., and Young, H. P. (1990): "Stochastic Evolutionary Game Dynamics," *Theoretical Population Biology* 38, 219-232.

Freidlin, M., and Wentzell, A. (1984): *Random Perturbations of Dynamical Systems.* New York: Springer Verlag.

Hahn, S. (1995): "The Long Run Equilibrium in an Asymmetric Coordination Game," mimeo, Harvard University.

Harsanyi, J., and Selten, R. (1988): *A General Theory of Equilibrium Selection in Games.* Cambridge: MIT Press.

Huck, S., and Oechssler, J. (1995): "The Indirect Approach to Explaining Fair Allocations,"

mimeo, Humboldt University.

Kandori, M., Mailath, G., and Rob, R. (1993): "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica* 61, 29-56.

Kandori, M., and Rob, R. (1992): "Evolution of Equilibria in the Long Run: A General Theory and Applications," CARESS Working Paper 92-06R.

Kandori, M., and Rob, R. (1993): "Bandwagon Effects and Long Run Technology Choice," University of Tokyo Discussion Paper 93-F-2.

Kim, Y. (1995): "Equilibrium Selection in n-Person Coordination Games," *Games and Economic Behavior*, forthcoming.

Maruta, T. (1995): "On the Relationship between Risk-Dominance and Stochastic Stability," mimeo, Kellogg-MEDS.

Morris, S., Rob, R., and Shin, H. (1995): "p-Dominance and Belief Potential," *Econometrica* 63, 145-157.

Nöldeke, G., and Samuelson, L. (1993): "An Evolutionary Analysis of Backward and Forward Induction," *Games and Economic Behavior* 5, 424-454.

Nöldeke, G., and Samuelson, L. (1994): "Learning to Signal in Markets," mimeo, University of Wisconsin.

Robson, A. and Vega Redondo, F. (1994): "Efficient Equilibrium Selection in Evolutionary Games with Random Matching," mimeo, University of Western Ontario.

Samuelson, L. (1994): "Stochastically Stable Sets in Games with Alternative Best Replies," *Journal of Economic Theory* 64, 35-65.

Young, H. P. (1993a): "The Evolution of Conventions," *Econometrica* 61, 57-84.

Young, H. P. (1993b): "An Evolutionary Model of Bargaining," *Journal of Economic Theory* 59, 145-168.